DOCUMENT RESUME

ED 263 967 JC 860 015

AUTHOR Yang, Shouu-Chyuan
TITLE The Single Sample Chi-Square Test: Lesson Plan.
INSTITUTION Hawaii Univ., Manoa. Western Curriculum Coordination Center.
PUB DATE 85
NOTE 24p.
PUB TYPE Guides - Classroom Use - Guides (For Teachers) (052)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Community Colleges; *Data Interpretation; Instructional Materials; Learning Modules; *Statistical Analysis; Teacher Developed Materials; Two Year Colleges
IDENTIFIERS *Chi Square Test

ABSTRACT

This lesson plan is designed to enable a student, after 50 minutes of instruction, to define the single sample chi-square test and explain the three conditions necessary to its proper use and the purpose of using it. The student will also be able to calculate and apply the single sample chi-square test in class according to the five steps with the given data. The lesson plan includes information on the course for which the lesson was designed; references; requirements for student materials, equipment, and audiovisual aids; lesson objectives; and special remarks for the instructor. Next, the plan presents a transcript of the informational content of the lesson, which provides time allocation suggestions and a key to the use of instructional methods and aids. The bulk of the lesson plan consists of handouts containing information, exercises, and assignments; and transparency masters. (EJV)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263967

THE SINGLE SAMPLE CHI-SQUARE TEST

Lesson Plan

Shouu-Chyuan Yang
Instructor

WESTERN CURRICULUM COORDINATION CENTER
University of Hawaii at Manoa

1985

JC 860 015

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

L. Zane

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

LESSON PLAN

LESSON PLAN NO. 1	NAME: Tahan Junior College	PAGE 1 OF 22 PAGES
COURSE OF INSTRUCTION: Chi-Square Test	TOTAL HOURS: 3	LESSON TITLE: The Single Sample Chi-Square Test
DAY AND DATE OF INSTRUCTION: Thursday, October 31, 1985	HOURS: 5:00-5:50 p.m. (50 min.)	PLACE: Kapiolani Community College, Diamond Head Campus, Bldg. Kauila, Room 204
NO. OF STUDENTS: 13 / REHEARSAL: Oct. 30, 9-11 a.m. Wist Annex 2-123	INSTRUCTOR AND ASSISTANTS: Shouu-chyuan Yang	STUDENT MATERIAL AND DRESS: Pencil or ball-point pen Calculator

REFERENCES:

Runyon, Richard P. and Haber, Audrey. Fundamentals of Behavioral Statistics. Random House, 1984, pp. 362-63.

Huck, Schuyler; Cornier, William; and Bounds, William G. Reading Statistics and Research, Harper and Row, 1974, pp. 216-18.

Theodorson, George A. and Theodorson, Achilles G. Modern Dictionary of Sociology, Apollo Editions, 1979.

EQUIPMENT AND AUDIO-VISUAL AIDS:

Overhead projector and screen
Transparencies
Chalk and eraser
Chalkboard
Pencil, paper and calculators (each student)

OBJECTIVES:

After 50 minutes of instruction, the student will be able to define the single sample chi-square test and explain the three conditions necessary to its proper use and the purpose of using it.

The student will be able to calculate and apply the single sample chi-square test in class according to the five steps with given data.

SPECIAL REMARKS:

This unit is part of a three (3) credits semester course in fundamental statistics.

1. Have overhead projector and screen ready.
2. Have transparencies ready.
3. Have chalkboard, chalk and eraser ready.

PRESENTATION

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS
1. Introduction During the last two weeks we discussed the different functions of parametric and non-parametric statistical procedures in testing hypotheses. Non-parametric statistical procedures are used in: 1) testing hypotheses that do	5 min	

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS
not specify normality or homogeneity of variance assumptions about the populations from which the samples were drawn (e.g., skewed distribution); 2) dealing with the nominal scale data (e.g., color: red, white; gender: male, female); and 3) dealing with small sample size (e.g., less than 30). We also discussed several non-parametric tests used in different instances, such as Spearman rank-order correlation coefficient (rho), Sign test, Wilcoxon matched-pairs signed-rank test, Median test, Mann-Whitney U test, Kruskal Wallis one-way ANOVA (analysis of variance) of ranks and Friedman two-way ANOVA of ranks. Today we are going to discuss another frequently used non-parametric test, the chi-square test. There are two kinds of chi-square tests: <u>single sample</u> chi-square test and <u>independent samples</u> chi-square test. We will only discuss the single sample chi-square test today.		
2. Explanation		
a. Definition of the chi-square test A statistical test used to determine the probability of differences between observed and theoretical, or expected, frequencies being due only to chance or being large enough to be statistically significant.	3 min	Lecture Handout #1 (2. a and b) Transparency #1 (same as Handout #1)
b. Function of the chi-square test To test whether or not a significant difference exists between the observed number of cases falling into each category and the expected number of cases, based on the null hypothesis.	4 min	

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS
Null hypothesis is a hypothesis of no differences or no relationship, that is, a hypothesis that states there is no significant differences between two or more groups in regard to a given variable or no relationship between the variables being studied. The X^2 test is one of the non-parametric tests; therefore, the conditions for using chi-square test are the same as using other non-parametric tests. They are: 1) testing hypotheses that do not specify normality or homogeneity of variance assumptions about the populations from which the samples are drawn; 2) dealing with the nominal scale data; and 3) dealing with small sample size.		
c. Example for calculating X^2 1) Suppose you are a market researcher hired by a soap manufacturer to conduct research on the packaging of its product. Realizing that color may be an important determinant of consumer selection, you conduct the following study. Six hundred homemakers, selected by an accepted sampling technique, are each given three differently packaged cakes of the same brand of soap. They are told that the three soaps are made according to different formulas and that the distinctive coloring of the packages is merely to aid their identification of each soap. One month later, you inform the homemakers that they are each to receive a free case of soap of their own choosing. Their selections are listed below.	20 min	Handout #2 [2. c. 1)] Transparency #2 (table in 2. c. 1) Handout #3 (chi-square formula) Handout #4 (Table A: critical ratio of chi-square) Transparency #3 (X^2 formula) Transparency #4 (Table A: two-tailed critical ratios of chi-square)

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS

	Color of Wrapper			Marginal
	Red	White	Brown	Total
No. of Homemakers Selecting	200	300	100	600

As shown in the table, the number of homemakers selecting in each color of wrapper are 200 for red, 300 for white, and 100 for brown. These are the observed number of cases falling into each category. We are to determine whether or not a significant difference exists between the observed number of cases falling into each category, and the expected number of cases, based on the null hypothesis.

2) Show the X^2 formula with transparency #2.

Question 1: What is the expected frequency in each category?

Answer: The null hypothesis would be that there is an equal preference for each color; therefore, 200 is the expected frequency in each category. $(\frac{600}{3} = 200)$

3) Insert the data into the X^2 formula:

Students do the calculation with their calculators.

Transparency #5

$$X^2 = \sum \frac{(fo - fe)^2}{fe}$$

$$= \frac{(200-200)^2}{200} + \frac{(300-200)^2}{200} + \frac{(100-200)^2}{200}$$

$$= 0 + 50.00 + 50.00$$

$$= 100$$

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS

To determine whether the X^2 value is significant enough to fall in the region of rejection, we have to find the value for critical region or regions of rejection. We have to find the degree of freedom by reducing one from the number of categories (n - 1), then find the critical value in the statistical tables under the particular df and level of significance (α). The level of significance can be set by the researcher according to the nature of the study. In this example:

n = 3
df = n - 1 = 2

Employing α = 0.01, we find that Table A indicates that a chi-square value of 9.21 or greater is required for significance. Since our obtained value of 100.00 is greater than 9.21, we may reject the null hypothesis and assert that color of packaging is a significant determinant for soap preference among homemakers.

On the other hand, if we have the results of homemakers selection like this:

Transparency #6

	Color of Wrapper			Marginal
	Red	White	Brown	Total
No. of Homemakers Selecting	210	220	170	600

OUTLINE OF INSTRUCTION TIME METHODS OF INSTRUCTION AND AIDS

then $X^2 = \frac{(210-200)^2}{200} + \frac{(220-200)^2}{200} + \frac{(170-200)^2}{200}$

$= \frac{100}{200} + \frac{400}{200} + \frac{900}{200}$

$= 0.5 + 2 + 4.5$

$= 7$

Because the X^2 value we obtained is smaller than 9.21 in $\alpha = .01$ we fail to reject the null hypothesis (that is, we accept the null hypothesis). Therefore, we conclude there is no difference between the observed frequencies and expected frequencies. That means the home-makers' selecting of soap was not affected by the color of wrapper.

3. Exercise 10 min Handout #5
(same as Handout #5) Transparency #7

Answer:

	Type A Cover	Type B Cover	Type C Cover	Type D Cover	Margin
Observed frequency	41	53	15	39	148
Expected frequency	37	37	37		

Expected frequency $= \frac{41+53+15+39}{4}$

$= \frac{148}{4}$

$= 37$

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS
$X^2 = \sum \frac{(fo - fe)^2}{fe}$ $= \frac{(41-37)^2 + (53-37)^2 + (15-37)^2 + (39-37)^2}{37}$ $= \frac{16 + 256 + 484 + 4}{37}$ $= \frac{760}{37}$ $= 20.54$ $df = N - 1$ $= 4 - 1 = 3$ Where $\alpha = .01$, critical value = 11.341. $X^2 = 20.54 > 11.341$ Therefore, the null hypothesis will be rejected. That is, the record buyers were affected by the type of album cover. The recording company is in the position to use Type B cover.		
4. Summary a. Definition of chi-square test Review Handout #1 b. Function of chi-square test Review Handout #1 Here we can see the conditions necessary for proper use of chi-square: Chi-square test is used when the researcher is --testing whether or not a significant difference exists between the observed number of cases falling into each category and the expected number of cases, based on the null hypothesis. --dealing with nominal scale or categorical data.	6 min	Handouts #1 and #3 Transparencies #1 and #3

OUTLINE OF INSTRUCTION	TIME	METHODS OF INSTRUCTION AND AIDS
—dealing with small sample size or testing hypotheses that do not specify normality or homogeneity of variance assumptions about the populations from which the sample was drawn.		
c. Calculation of chi-square test Review Handout 3		
d. Questions for review		
Q1. What is the chi-square test? A. A statistical test used to determine the probability of differences between observed and expected frequencies being due only to chance or being large enough to be statistically significant.		
Q2. What is the observed frequency or "fo"? A. The observed or actual number of cases falling into each category.		
Q3. What is the expected frequency or "fe"? A. The expected number of cases in each category.		
Q4. What is Σ ? A. Sum of.		
Q5. What is "n"? A. The number of categories.		
Q6. What is "alpha" (α)? A. The level of significance.		
Q7. What is "df"? A. The degree of freedom which can be obtained by subtracting one from the number of categories, i.e., n - 1.		
Q8. If we have n = 6, what is the critical ratio of α = .01? A. n = 6 df = 6 - 1 = 5 When α = .01, the critical ratio is 15.086.		
5. Homework	2 min	Handout #6 (homework)

DEFINITION OF THE CHI-SQUARE TEST:

A statistical test used to determine the probability of differences between observed and theoretical, or expected, frequencies being due only to chance or being large enough to be statistically significant.

FUNCTION OF THE CHI-SQUARE TEST:

To test whether or not a significant difference exists between the observed number of cases falling into each category, and the expected number of cases, based on the null hypothesis.

Suppose you are a market researcher hired by a soap manufacturer to conduct research on the packaging of its product. Realizing that color may be an important determinant of consumer selection, you conduct the following study. Six hundred homemakers, selected by an accepted sampling technique, are each given three differently packaged cakes of the same brand of soap. They are told that the three soaps are made according to different formulas and that the distinctive coloring of the packages is merely to aid their identification of each soap. One month later, you inform the homemakers that they are each to receive a free case of soap of their own choosing. Their selections are listed below.

	COLOR OF WRAPPER			MARGINAL
	Red	White	Brown	TOTAL
NUMBER OF HOMEMAKERS SELECTING	200	300	100	600

CHI-SQUARE FORMULA OF ONE-VARIABLE TEST:

$$X^2 = \sum_i \frac{(\text{Observed frequency}_i - \text{Expected frequency}_i)^2}{\text{Expected frequency}_i}$$

OR

$$X^2 = \sum_i \frac{(fo_i - fe_i)^2}{fe_i}$$

Where $\sum$ = sum of

fo = observed frequency

fe = expected frequency

i = number of each category

CALCULATION OF THE CHI-SQUARE ONE-VARIABLE TEST AND THE APPLICATION:

STEP 1. Calculate expected frequency by dividing total number of frequencies falling in all categories with the number of categories.

That is, Expected frequency = $\frac{\text{Total frequency}}{\text{Number of categories}}$

OR $fe = \frac{\text{Total}}{N}$

STEP 2. Insert number obtained into chi-square formula.

$$X^2 = \sum_i \frac{(fo_i - fe_i)^2}{fe_i}$$

STEP 3. Calculate df = N - 1.

STEP 4. Find critical values for various (α) levels of significance from the statistical tables (in this course, Table A).

STEP 5. Compare the chi-square value obtained and the critical value to see whether the null hypothesis is rejected or not, to determine if a significant difference exists between the observed frequency in each category and the expected frequency.

Table A
Two-tailed critical ratios of χ^2.

Degrees of freedom df	.10	.05	.02	.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.341
4	7.779	9.488	11.668	13.277
5	9.236	11.070	13.388	15.086
6	10.645	12.592	15.033	16.812
7	12.017	14.067	16.622	18.475
8	13.362	15.507	18.168	20.090
9	14.684	16.919	19.679	21.666
10	15.987	18.307	21.161	23.209
11	17.275	19.675	22.618	24.725
12	18.549	21.026	24.054	26.217
13	19.812	22.362	25.472	27.688
14	21.064	23.685	26.873	29.141
15	22.307	24.996	28.259	30.578
16	23.542	26.296	29.633	32.000
17	24.769	27.587	30.995	33.409
18	25.989	28.869	32.346	34.805
19	27.204	30.144	33.687	36.191
20	28.412	31.410	35.020	37.566
21	29.615	32.671	36.343	38.932
22	30.813	33.924	37.659	40.289
23	32.007	35.172	38.968	41.638
24	33.196	36.415	40.270	42.980
25	34.382	37.652	41.566	44.314
26	35.563	38.885	42.856	45.642
27	36.741	40.113	44.140	46.963
28	37.916	41.337	45.419	48.278
29	39.087	42.557	46.693	49.588
30	40.256	43.773	47.962	50.892

BEST COPY AVAILABLE

EXERCISE:

Suppose that a recording company is interested in the type of cover to put on a record album. It sends the same record with four different covers to stores. At the end of a month, it is found that the following number of albums have been sold.

Type A Cover	Type B Cover	Type C Cover	Type D Cover
41	53	15	39

Is the recording company in a position to determine which type of cover it should use?

HOMEWORK:

Answer the following questions using the chi-square one-variable test.

1. A coin is tossed 20 times and lands heads up 13 times. Test whether the coin is biased.

2. Calculate X^2 for the following table and see whether there is a significant difference between observed number and expected number.

Observed	6	8	4	10
Expected	7	9	6	6

3. A model predicts that the grades on an achievement test will be as described in the following table. Observed grades are also given.

Grade	A	B	C	D	E
Observed grade	30	40	60	40	30
Expected grade	40	40	40	40	40

Use X^2 to test the validity of the model.

HOMEWORK ANSWERS

1.

Coin tossing	Head	Tail	Margin
Observed numbers	13	7	20
Expected numbers	10	10	

$$X^2 = \frac{(13 - 10)^2 + (7 - 10)^2}{10}$$

$$= \frac{9 + 9}{10} = 1.8$$

$df = 2 - 1 = 1$

When $\alpha = .01$, critical $X^2 = 6.635$

$X^2 = 1.8 < 6.635$
Therefore, the null hypothesis is accepted. The coin is not biased.

2. $$X^2 = \sum \frac{(6-7)^2}{7} + \frac{(8-9)^2}{9} + \frac{(4-6)^2}{6} + \frac{(10-6)^2}{6}$$

$$= \frac{1}{7} + \frac{1}{9} + \frac{4}{6} + \frac{16}{6}$$

$$= \frac{54 + 42 + 252 + 1008}{378}$$

$$= \frac{1356}{378} = 3.587$$

3. $$X^2 = \sum \frac{(30-40)^2 + (40-40)^2 + (60-40)^2 + (40-40)^2 + (30-40)^2}{40}$$

$$= \frac{100 + 0 + 400 + 0 + 100}{40}$$

$$= \frac{600}{40} = 15$$

$df = 5 - 1 = 4$
When $\alpha = .01$, the critical $X^2 = 13.277$. $X^2 = 15 > 13.277$
So, we reject the null hypothesis. The model was not validated.

DEFINITION OF THE CHI-SQUARE TEST:

A statistical test used to determine the probability of differences between observed and theoretical, or expected, frequencies being due only to chance or being large enough to be statistically significant.

FUNCTION OF THE CHI-SQUARE TEST:

To test whether or not a significant difference exists between the observed number of cases falling into each category, and the expected number of cases, based on the null hypothesis.

Suppose you are a market researcher hired by a soap manufacturer to conduct research on the packaging of its product. Realizing that color may be an important determinant of consumer selection, you conduct the following study. Six hundred homemakers, selected by an accepted sampling technique, are each given three differently packaged cakes of the same brand of soap. They are told that the three soaps are made according to different formulas and that the distinctive coloring of the packages is merely to aid their identification of each soap. One month later, you inform the homemakers that they are each to receive a free case of soap of their own choosing. Their selections are listed below.

	COLOR OF WRAPPER			MARGINAL
	Red	White	Brown	TOTAL
NUMBER OF HOMEMAKERS SELECTING	200	300	100	

CHI-SQUARE FORMULA OF ONE-VARIABLE TEST:

$$X^2 = \sum_i \frac{(\text{Observed frequency}_i - \text{Expected frequency}_i)^2}{\text{Expected frequency}_i} \quad \text{OR}$$

$$X^2 = \sum_i \frac{(fo_i - fe_i)^2}{fe_i}$$

Where $\sum$ = sum of
fo = observed frequency
fe = expected frequency
i = number of each category

CALCULATION OF THE CHI-SQUARE ONE-VARIABLE TEST AND THE APPLICATION:

STEP 1. Calculate expected frequency by dividing total number of frequencies falling in all categories with the number of categories.

That is, Expected frequency = $\frac{\text{Total frequency}}{\text{Number of categories}}$

OR $fe = \frac{\text{Total}}{N}$

STEP 2. Insert number obtained into chi-square formula.

$$X^2 = \sum_i \frac{(fo_i - fe_i)^2}{fe_i}$$

STEP 3. Calculate df = N - 1.

STEP 4. Find critical values for various α levels of significance from the statistical tables (in this course, Table A).

STEP 5. Compare the chi-square value obtained and the critical value to see whether the null hypothesis is rejected or not, to determine if a significant difference exists between the observed frequency in each category and the expected frequency.

TABLE A
TWO-TAILED CRITICAL RATIOS OF X2

Degrees of freedom dF	.10	.05	.02	.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.341
4	7.779	9.488	11.668	13.277
5	9.236	11.070	13.388	15.086
6	10.645	12.592	15.033	16.812
7	12.017	14.067	16.622	18.475
8	13.362	15.507	18.168	20.090
9	14.684	16.919	19.679	21.666
10	15.987	18.307	21.161	23.209
11	17.275	19.675	22.618	24.725
12	18.549	21.026	24.054	26.217
13	19.812	22.362	25.472	27.688
14	21.064	23.685	26.873	29.141
15	22.307	24.996	28.259	30.578
16	23.542	26.296	29.633	32.000
17	24.769	27.587	30.995	33.409
18	25.989	28.869	32.346	34.805
19	27.204	30.144	33.687	36.191
20	28.412	31.410	35.020	37.566
21	29.615	32.671	36.343	38.932
22	30.813	33.924	37.659	40.289
23	32.007	35.172	38.968	41.638
24	33.196	36.415	40.270	42.980
25	34.382	37.652	41.566	44.314
26	35.563	38.885	42.856	45.642
27	36.741	40.113	44.140	46.963
28	37.916	41.337	45.419	48.278
29	39.087	42.557	46.693	49.588
30	40.256	43.773	47.962	50.892

	COLOR OF WRAPPER			MARGINAL TOTALS
	RED	WHITE	BROWN	
NO. OF HOMEMAKERS SELECTING	210	220	170	600
NO. EXPECTED	200	200	200	

$$\chi^2 = \frac{(210-200)^2}{200} + \frac{(220-200)^2}{200} + \frac{(170-200)^2}{200}$$

$$= \frac{100}{200} + \frac{400}{200} + \frac{900}{200}$$

$$= .5 + 2 + 4.5$$

$$= 7$$

$df = 2$

$\alpha = .01$

CRITICAL VALUE = 9.21

FAIL TO REJECT (THAT IS, TO ACCEPT) THE NULL HYPOTHESIS.

$$f_e = \frac{\text{TOTAL}}{n}$$

$$= \frac{600}{3}$$

$$= 200$$

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}$$

$$= \frac{(200-200)^2}{200} + \frac{(300-200)^2}{200} + \frac{(100-200)^2}{200}$$

$$= 0 + 50 + 50$$

$$= 100$$

$n = 3$

$df = n - 1 = 2$

$\alpha = .01$

CRITICAL VALUE = 9.21

	TYPE A COVER	TYPE B COVER	TYPE C COVER	TYPE D COVER	TOTAL
OBSERVED FREQUENCY	41	53	15	39	
EXPECTED FREQUENCY					

$$f_e = \frac{\Sigma f_o}{N} = \frac{41+53+15+39}{3} = \frac{148}{3} = 37$$

$$\chi^2 = \sum \frac{(f_{o_i} - f_{e_i})^2}{f_{e_i}}$$

$$= \frac{(41-37)^2}{37} + \frac{(53-37)^2}{37} + \frac{(15-37)^2}{37} + \frac{(39-37)^2}{37}$$

$$= \frac{16+256+484+4}{37}$$

$$= \frac{760}{37} = 20.54$$

$df = N-1 = 4-1 = 3$

$\alpha = .01$ critical value = 11.341

$\chi^2 = 20.54 > 11.34$ reject null hypothesis

ERIC CLEARINGHOUSE FOR JUNIOR COLLEGES
JAN 31 1986